

04/13/2001

#3  
OIPE

## RAW SEQUENCE LISTING

DATE: 04/09/2001

PATENT APPLICATION: US/09/804,690

TIME: 11:59:35

Input Set : N:\Crf3\RULE60\09804690.txt

Output Set: N:\CRF3\04092001\I804690.raw

## SEQUENCE LISTING

## 5 (1) GENERAL INFORMATION:

7 (i) APPLICANT: LI, Limin

8 COHEN, Stanley N

11 (ii) TITLE OF INVENTION: MAMMALIAN TUMOR SUSCEPTIBILITY GENES AND  
12 THEIR USES

14 (iii) NUMBER OF SEQUENCES: 20

16 (iv) CORRESPONDENCE ADDRESS:

17 (A) ADDRESSEE: FISH AND RICHARDSON, P.C.

18 (B) STREET: 2200 SAND HILL ROAD

19 (C) CITY: MENLO PARK

20 (D) STATE: CA

21 (E) COUNTRY: USA

22 (F) ZIP: 94025

25 (v) COMPUTER READABLE FORM:

26 (A) MEDIUM TYPE: Floppy disk

27 (B) COMPUTER: IBM PC compatible

28 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

29 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

31 (vi) CURRENT APPLICATION DATA:

-C--&gt; 32 (A) APPLICATION NUMBER: US/09/804,690

C--&gt; 33 (B) FILING DATE: 12-Mar-2001

34 (C) CLASSIFICATION:

36 (vii) PRIOR APPLICATION DATA:

37 (A) APPLICATION NUMBER: 09/146,187

38 (B) FILING DATE:

40 (viii) ATTORNEY/AGENT INFORMATION:

41 (A) NAME: SHERWOOD, Pamela J.

42 (B) REGISTRATION NUMBER: 36,677

44 (ix) TELECOMMUNICATION INFORMATION:

45 (A) TELEPHONE: 415-781-1989

46 (B) TELEFAX: 415-398-3249

47 (C) TELEX: 910 277299

50 (2) INFORMATION FOR SEQ ID NO: 1:

52 (i) SEQUENCE CHARACTERISTICS:

53 (A) LENGTH: 1448 base pairs

54 (B) TYPE: nucleic acid

55 (C) STRANDEDNESS: double

56 (D) TOPOLOGY: linear

58 (ii) MOLECULE TYPE: cDNA

61 (ix) FEATURE:

62 (A) NAME/KEY: CDS

63 (B) LOCATION: 61..1203

66 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

68 CCCCTCTGCC TGTGGGACG GAGGAGCGCG CCATGGCTGT CCGAGAGTCA GCTGAAGAAG 60

70 ATG ATG TCC AAG TAC AAA TAT AGA GAT CTA ACC GTC CGT CAA ACT GTC 108

71 Met Met Ser Lys Tyr Lys Tyr Arg Asp Leu Thr Val Arg Gln Thr Val

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/804,690

DATE: 04/09/2001

TIME: 11:59:35

Input Set : N:\Crf3\RULE60\09804690.txt

Output Set: N:\CRF3\04092001\I804690.raw

72	1	5	10	15	
74	AAT GTC ATC GCT ATG TAC AAA GAT CTC AAA CCT GTA TTG GAT TCA TAT	156			
75	Asn Val Ile Ala Met Tyr Lys Asp Leu Lys Pro Val Leu Asp Ser Tyr				
76	20 25 30				
78	GTT TTT AAT GAT GGC AGT TCC AGG GAG CTG GTG AAC CTC ACT GGT ACA	204			
79	Val Phe Asn Asp Gly Ser Ser Arg Glu Leu Val Asn Leu Thr Gly Thr				
80	35 40 45				
82	ATC CCA GTG CGT TAT CGA GGT AAT ATA TAT AAT ATT CCA ATA TGC CTG	252			
83	Ile Pro Val Arg Tyr Arg Gly Asn Ile Tyr Asn Ile Pro Ile Cys Leu				
84	50 55 60				
86	TGG CTG CTG GAC ACA TAC CCA TAT AAC CCC CCT ATC TGT TTT GTT AAG	300			
87	Trp Leu Leu Asp Thr Tyr Pro Tyr Asn Pro Pro Ile Cys Phe Val Lys				
88	65 70 75 80				
90	CCT ACT AGT TCA ATG ACT ATT AAA ACA GGA AAG CAT GTG GAT GCA AAT	348			
91	Pro Thr Ser Ser Met Thr Ile Lys Thr Gly Lys His Val Asp Ala Asn				
92	85 90 95				
94	GGG AAA ATC TAC CTA CCT TAT CTA CAT GAC TGG AAA CAT CCA CGG TCA	396			
95	Gly Lys Ile Tyr Leu Pro Tyr Leu His Asp Trp Lys His Pro Arg Ser				
96	100 105 110				
98	GAG TTG CTG GAG CTT ATT CAA ATC ATG ATT GTG ATA TTT GGA GAG GAG	444			
99	Glu Leu Leu Glu Leu Ile Gln Ile Met Ile Val Ile Phe Gly Glu Glu				
100	115 120 125				
102	CCT CCA GTG TTC TCC CGG CCT ACT GTT TCT GCA TCC TAC CCA CCA TAC	492			
103	Pro Pro Val Phe Ser Arg Pro Thr Val Ser Ala Ser Tyr Pro Pro Tyr				
104	130 135 140				
106	ACA GCA ACA GGG CCA CCA AAT ACC TCC TAC ATG CCA GGC ATG CCA AGT	540			
107	Thr Ala Thr Gly Pro Pro Asn Thr Ser Tyr Met Pro Gly Met Pro Ser				
108	145 150 155 160				
110	GGA ATC TCT GCA TAT CCA TCT GGA TAC CCT CCC AAC CCC AGT GGT TAT	588			
111	Gly Ile Ser Ala Tyr Pro Ser Gly Tyr Pro Pro Asn Pro Ser Gly Tyr				
112	165 170 175				
114	CCT GGC TGT CCT TAC CCA CCT GCT GGC CCA TAC CCT GCC ACA ACA AGC	636			
115	Pro Gly Cys Pro Tyr Pro Pro Ala Gly Pro Tyr Pro Ala Thr Thr Ser				
116	180 185 190				
118	TCA CAG TAC CCT TCC CAG CCT CCT GTG ACC ACT GTT GGT CCC AGC AGA	684			
119	Ser Gln Tyr Pro Ser Gln Pro Pro Val Thr Thr Val Gly Pro Ser Arg				
120	195 200 205				
122	GAT GGC ACA ATC AGT GAG GAC ACT ATC CGT GCA TCT CTC ATC TCA GCA	732			
123	Asp Gly Thr Ile Ser Glu Asp Thr Ile Arg Ala Ser Leu Ile Ser Ala				
124	210 215 220				
126	GTC AGT GAC AAA CTG AGA TGG CGG ATG AAG GAG GAA ATG GAT GGT GCC	780			
127	Val Ser Asp Lys Leu Arg Trp Arg Met Lys Glu Glu Met Asp Gly Ala				
128	225 230 235 240				
130	CAG GCA GAG CTT AAT GCC TTG AAA CGA ACA GAG GAA GAT CTG AAA AAA	828			
131	Gln Ala Glu Leu Asn Ala Leu Lys Arg Thr Glu Glu Asp Leu Lys Lys				
132	245 250 255				
134	GGC CAC CAG AAA CTG GAA GAG ATG GTC ACC CGC TTA GAT CAA GAA GTA	876			
135	Gly His Gln Lys Leu Glu Glu Met Val Thr Arg Leu Asp Gln Glu Val				
136	260 265 270				

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/804,690

DATE: 04/09/2001

TIME: 11:59:35

Input Set : N:\Crif3\RULE60\09804690.txt

Output Set: N:\CRF3\04092001\I804690.raw

```

138 GCT GAA GTT GAT AAA AAC ATA GAA CTT TTG AAA AAG AAG GAT GAA GAA      924
139 Ala Glu Val Asp Lys Asn Ile Glu Leu Leu Lys Lys Lys Asp Glu Glu
140      275      280      285
142 CTA AGT TCT GCT CTG GAG AAA ATG GAA AAT CAA TCT GAA AAT AAT GAT      972
143 Leu Ser Ser Ala Leu Glu Lys Met Glu Asn Gln Ser Glu Asn Asn Asp
144      290      295      300
146 ATT GAT GAA GTT ATC ATT CCC ACA GCC CCA CTG TAT AAA CAG ATT CTA      1020
147 Ile Asp Glu Val Ile Ile Pro Thr Ala Pro Leu Tyr Lys Gln Ile Leu
148 305      310      315      320
150 AAT CTG TAT GCA GAG GAA AAT GCT ATT GAA GAC ACT ATC TTT TAC CTT      1068
151 Asn Leu Tyr Ala Glu Glu Asn Ala Ile Glu Asp Thr Ile Phe Tyr Leu
152      325      330      335
154 GGA GAA GCT TTG CGG CGG GGA GTC ATA GAC CTG GAT GTG TTC CTG AAA      1116
155 Gly Glu Ala Leu Arg Arg Gly Val Ile Asp Leu Asp Val Phe Leu Lys
156      340      345      350
158 CAC GTC CGC CTC CTG TCC CGT AAA CAG TTC CAG CTA AGG GCA CTA ATG      1164
159 His Val Arg Leu Leu Ser Arg Lys Gln Phe Gln Leu Arg Ala Leu Met
160      355      360      365
162 CAA AAG GCA AGG AAG ACT GCG GGC CTT AGT GAC CTC TAC TGACATGTGC      1213
163 Gln Lys Ala Arg Lys Thr Ala Gly Leu Ser Asp Leu Tyr
164      370      375      380
166 TGTCAGCTGG AGACCGACCT CTCCGTAAAG CATTCTTTTC TTCTTCTTTT TCTCATCAGT      1273
168 AGAACCCACA ATAAGTTATT GCAGTTTATC ATTCAAGTGT TAAATATTTT GAATCAATAA      1333
170 TATATTTTCT GTTTCCTTTG GGTA AAAACT GGCTTTTATT AATGCACTTT CTACCCTCTG      1393
172 TAAGCGTCTG TGCTGTGCTG GGA CTGACTG GGCTAAATAA AATTTGTTGC ATAAA      1448
175 (2) INFORMATION FOR SEQ ID NO: 2:
177 (i) SEQUENCE CHARACTERISTICS:
178 (A) LENGTH: 381 amino acids
179 (B) TYPE: amino acid
180 (D) TOPOLOGY: linear
182 (ii) MOLECULE TYPE: protein
184 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
186 Met Met Ser Lys Tyr Lys Tyr Arg Asp Leu Thr Val Arg Gln Thr Val
187 1 5 10 15
189 Asn Val Ile Ala Met Tyr Lys Asp Leu Lys Pro Val Leu Asp Ser Tyr
190 20 25 30
192 Val Phe Asn Asp Gly Ser Ser Arg Glu Leu Val Asn Leu Thr Gly Thr
193 35 40 45
195 Ile Pro Val Arg Tyr Arg Gly Asn Ile Tyr Asn Ile Pro Ile Cys Leu
196 50 55 60
198 Trp Leu Leu Asp Thr Tyr Pro Tyr Asn Pro Pro Ile Cys Phe Val Lys
199 65 70 75 80
201 Pro Thr Ser Ser Met Thr Ile Lys Thr Gly Lys His Val Asp Ala Asn
202 85 90 95
204 Gly Lys Ile Tyr Leu Pro Tyr Leu His Asp Trp Lys His Pro Arg Ser
205 100 105 110
207 Glu Leu Leu Glu Leu Ile Gln Ile Met Ile Val Ile Phe Gly Glu Glu
208 115 120 125
210 Pro Pro Val Phe Ser Arg Pro Thr Val Ser Ala Ser Tyr Pro Pro Tyr

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/804,690

DATE: 04/09/2001

TIME: 11:59:35

Input Set : N:\Crf3\RULE60\09804690.txt

Output Set: N:\CRF3\04092001\I804690.raw

```

211      130      135      140
213 Thr Ala Thr Gly Pro Pro Asn Thr Ser Tyr Met Pro Gly Met Pro Ser
214 145      150      155      160
216 Gly Ile Ser Ala Tyr Pro Ser Gly Tyr Pro Pro Asn Pro Ser Gly Tyr
217      165      170      175
219 Pro Gly Cys Pro Tyr Pro Pro Ala Gly Pro Tyr Pro Ala Thr Thr Ser
220      180      185      190
222 Ser Gln Tyr Pro Ser Gln Pro Pro Val Thr Thr Val Gly Pro Ser Arg
223      195      200      205
225 Asp Gly Thr Ile Ser Glu Asp Thr Ile Arg Ala Ser Leu Ile Ser Ala
226      210      215      220
228 Val Ser Asp Lys Leu Arg Trp Arg Met Lys Glu Glu Met Asp Gly Ala
229 225      230      235      240
231 Gln Ala Glu Leu Asn Ala Leu Lys Arg Thr Glu Glu Asp Leu Lys Lys
232      245      250      255
234 Gly His Gln Lys Leu Glu Glu Met Val Thr Arg Leu Asp Gln Glu Val
235      260      265      270
237 Ala Glu Val Asp Lys Asn Ile Glu Leu Leu Lys Lys Lys Asp Glu Glu
238      275      280      285
240 Leu Ser Ser Ala Leu Glu Lys Met Glu Asn Gln Ser Glu Asn Asn Asp
241      290      295      300
243 Ile Asp Glu Val Ile Ile Pro Thr Ala Pro Leu Tyr Lys Gln Ile Leu
244 305      310      315      320
246 Asn Leu Tyr Ala Glu Glu Asn Ala Ile Glu Asp Thr Ile Phe Tyr Leu
247      325      330      335
249 Gly Glu Ala Leu Arg Arg Gly Val Ile Asp Leu Asp Val Phe Leu Lys
250      340      345      350
252 His Val Arg Leu Leu Ser Arg Lys Gln Phe Gln Leu Arg Ala Leu Met
253      355      360      365
255 Gln Lys Ala Arg Lys Thr Ala Gly Leu Ser Asp Leu Tyr
256      370      375      380
258 (2) INFORMATION FOR SEQ ID NO: 3:
260 (i) SEQUENCE CHARACTERISTICS:
261 (A) LENGTH: 1494 base pairs
262 (B) TYPE: nucleic acid
263 (C) STRANDEDNESS: double
264 (D) TOPOLOGY: linear
266 (ii) MOLECULE TYPE: cDNA
269 (ix) FEATURE:
270 (A) NAME/KEY: CDS
271 (B) LOCATION: 120..1259
274 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
276 GAAGGGGGTG TGCGATTGTG TGGGACGGTC TGGGGCAGCC ACAGCGGCTG ACCNCNTNGC 60
278 CTGCGGGGAA GGGAGTCGCC AGGGCCCCGC ATCGGGTGTC GGAGAGCCAG CTCAAGAAAA 120
280 TGGTGTCCAA GTACAAATAC AGAGACCTAA CTGTACGTGA AACTGTCAAT GTTATTACTC 180
282 TATACAAAGA TCTCAAACCT GTTTTGGATT CATATGTTTT TAACGATGGC AGTTCCAGGG 240
284 AACTAATGAA CCTCACTGGA ACAATCCCTG TGCCTTATAG AGGTAATACA TACAATATTC 300
286 CAATATGCCT ATGGCTACTG GACACATACC CATATAATCC CCCTATCTGT TTTGTTAAGC 360
288 CTA TAGTTC AATGACTATT AAAACAGGAA AGCATGTTGA TGCAAATGGG AAGATATATC 420

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/804,690

DATE: 04/09/2001

TIME: 11:59:35

Input Set : N:\Crif3\RULE60\09804690.txt

Output Set: N:\CRF3\04092001\I804690.raw

```

290 TTCCTTATCT ACATGAATGG AAACACCCAC AGTCAGACTT GTTGGGGCTT ATTCAGGTCA 480
292 TGATTGTGGT ATTTGGAGAT GAACCTCCAG TCTTCTCTCG TCCTATTTCG GCATCCTATC 540
294 CGCCATACCA GGCAACGGGG CCACCAAATA CTTCTACAT GCCAGGCATG CCAGGTGGAA 600
296 TCTCTCCATA CCCATCCGGA TACCCTCCCA ATCCCAGTGG TTACCCAGGC TGTCTTACC 660
298 CACCTGGTGG TCCATATCCT GCCACAACAA GTTCTCAGTA CCCTTCTCAG CCTCCTGTGA 720
300 CCACTGTTGG TCCCAGTAGG GATGGCACAA TCAGCGAGGA CACCATCCGA GCCTCTCTCA 780
302 TCTCTGCGGT CAGTGACAAA CTGAGATGGC GGATGAAGGA GGAAATGGAT CGTGCCAGG 840
304 CAGAGCTCAA TGCTTGAAA CGAACAGAAG AAGACCTGAA AAAGGGTCAC CAGAAACTGG 900
306 AAGAGATGGT TACCCGTTTA GATCAAGAAG TAGCCGAGGT TGATAAAAC ATAGAACTTT 960
308 TGAAAAAGAA GGATGAAGAA CTCAGTTCTG CTCTGGAAAA AATGGAAAT CAGTCTGAAA 1020
310 ACAATGATAT CGATGAAGTT ATCATTCCCA CAGCTCCCTT ATACAAACAG ATCCTGAATC 1080
312 TGTATGCAGA AGAAAACGCT ATTGAAGACA CTATCTTTTA CTTGGGAGAA GCCTTGAGAA 1140
314 GGGGCGTGAT AGACCTGGAT GTCTTCCTGA AGCATGTACG TCTTCTGTCC CGTAAACAGT 1200
316 TCCAGCTGAG GGCACAAATG CAAAAAGCAA GAAAGACTGC CGGTCTCAGT GACCTCTACT 1260
318 GACTTCTCTG ATACCAGCTG GAGGTTGAGC TCTTCTTAAA GTATTCTTCT CTTCCTTTTA 1320
320 TCAGTAGGTG CCCAGAATAA GTTATTGCAG TTTATCATTC AAGTGTAATA TATTTTGAAT 1380
322 CAATAATATA TTTTCTGTTT TCTTTTGTA AAGACTGGCT TTTATTAATG CACTTTCTAT 1440
324 CCTCTGTAAT CTTTTGTGC TGAATGTTGG GACTGCTAAA TAAATTTGT TTTT 1494

```

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 380 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

337 Met Val Ser Lys Tyr Lys Tyr Arg Asp Leu Thr Val Arg Glu Thr Val
338 1 5 10 15
340 Asn Val Ile Thr Leu Tyr Lys Asp Leu Lys Pro Val Leu Asp Ser Tyr
341 20 25 30
343 Val Phe Asn Asp Gly Ser Ser Arg Glu Leu Met Asn Leu Thr Gly Thr
344 35 40 45
346 Ile Pro Val Pro Tyr Arg Gly Asn Thr Tyr Asn Ile Pro Ile Cys Leu
347 50 55 60
349 Trp Leu Leu Asp Thr Tyr Pro Tyr Asn Pro Pro Ile Cys Phe Val Lys
350 65 70 75 80
352 Pro Thr Ser Ser Met Thr Ile Lys Thr Gly Lys His Val Asp Ala Asn
353 85 90 95
355 Gly Lys Ile Tyr Leu Pro Tyr Leu His Glu Trp Lys His Pro Gln Ser
356 100 105 110
358 Asp Leu Leu Gly Leu Ile Gln Val Met Ile Val Val Phe Gly Asp Glu
359 115 120 125
361 Pro Pro Val Phe Ser Arg Pro Ile Ser Ala Ser Tyr Pro Pro Tyr Gln
362 130 135 140
364 Ala Thr Gly Pro Pro Asn Thr Ser Tyr Met Pro Gly Met Pro Gly Gly
365 145 150 155 160
367 Ile Ser Pro Tyr Pro Ser Gly Tyr Pro Pro Asn Pro Ser Gly Tyr Pro
368 165 170 175
370 Gly Cys Pro Tyr Pro Pro Gly Gly Pro Tyr Pro Ala Thr Thr Ser Ser
371 180 185 190

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/804,690

DATE: 04/09/2001

TIME: 11:59:36

Input Set : N:\Crf3\RULE60\09804690.txt

Output Set: N:\CRF3\04092001\I804690.raw

L:32 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:33 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:442 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6